# 3. Enhancing Instructional Programming and Student Achievement with Curriculum-Based Measurement

Lynn S. Fuchs
*George Peabody College of Vanderbilt University*, Lynn.Fuchs@Vanderbilt.Edu

3

# Enhancing Instructional Programming and Student Achievement with Curriculum-Based Measurement

Lynn S. Fuchs

*George Peabody College of Vanderbilt University*

Curriculum-based measurement (CBM) is a form of curriculum-based assessment. As such, CBM has three features in common with all curriculum-based assessment approaches (Tucker, 1987): Test stimuli are drawn from the student's curriculum; assessment is ongoing and repeated across time; and assessment data are used to formulate instructional decisions.

Despite these similarities to other forms of curriculum-based assessment, CBM is distinctive because of two important features: It measures student proficiency across the annual curriculum and relies on standardized, prescriptive measurement methods (Fuchs & Deno, in press). The first purpose of this chapter is to explain these two features of CBM, by contrasting the CBM model to the predominant, mastery measurement form of curriculum-based assessment.

The second objective of this paper is to demonstrate how CBM databases can be used to help formulate instructional decisions. Within this context, research investigating the efficacy of each instructional use is reviewed.

## THE CURRICULUM-BASED MEASUREMENT MODEL

As indicated above, two important features of curriculum-based measurement (CBM) are (a) its focus on measuring student proficiency across the annual curriculum and (b) its use of a standardized, prescriptive measurement methodology, with demonstrated psychometric acceptability. To explain each of these features, I contrast CBM to the more common, predominant form of curriculum-based assessment known as mastery measurement. Within this section, I first explain and provide an example of mastery measurement. Then, I explain CBM and provide an example. Finally, the salient differences between mastery measurement and CBM are explored.

### Mastery Measurement

Mastery measurement is the most common form of curriculum-based assessment (see Shinn, Rosenfield, & Knutson, 1989 for discussion of different types of curriculum-based assessment). Mastery measurement describes student mastery of a series of short-term instructional objectives or instructional levels (see Blankenship, 1985 and Gickling & Thompson, 1985 for explanation of these forms of mastery measurement). So, for example, let us say that Mrs. P. wants Dolly to master the fourth-grade computation curriculum. That is, by June Mrs. P. wants Dolly to compute accurately all problem types encompassed within the fourth-grade curriculum. In designing a mastery measurement system, Mrs. P. would begin by completing two large tasks. She would (a) determine a sensible instructional sequence for the fourth-grade computation curriculum and (b) design a criterion-referenced testing procedure to match each step in that instructional sequence.

Let us say, for example, that after careful inspection of the fourth-grade computation curriculum, Mrs. P. identified the skills listed in Table 1. These are the universe of problem types incorporated within her fourth-grade curriculum. She further determined that a logical sequence of skills for instruction were the following: multidigit addition with regrouping, multidigit subtraction with regrouping, multiplication facts (factors to 9), division facts (divisors 6-9), multiplying two 2-digit numbers without regrouping, multiplying 1- or 2-digit numbers with regrouping, dividing 3- by 1-digit numbers without remainders, dividing 2- or 3- by 1-digit numbers with remainders, adding and subtracting mixed decimals to hundredths, and adding and subtracting simple or mixed fractions without regrouping.

Table 1
Fourth Grade Curriculum

| Sequence | Skill | Proportion |
|---|---|---|
| 1 | Multidigit addition with regrouping | 12% |
| 2 | Multidigit subtraction with regrouping | 4% |
| 3 | Multiplication facts, factors to 9 | 24% |
| 4 | Division facts, divisors 6-9 | 16% |
| 5 | Multiplying two 2-digit numbers, no regrouping | 4% |
| 6 | Multiplying 1- or 2-digit numbers,with regrouping | 12% |
| 7 | Dividing 3- by 1-digit numbers, no remainder | 4% |
| 8 | Dividing 2- or 3- by 1-digit numbers, with remainder | 4% |
| 9 | Adding and subtracting mixed decimals to hundredths | 8% |
| 10 | Adding and subtracting simple or mixed fractions, no regrouping | 12% |

Having established the instructional sequence, Mrs. P.'s second major task in establishing a mastery measurement system would be to design a criterion-referenced testing procedure for each step in her instructional hierarchy. By definition, Mrs. P. would begin by measuring the first skill in the sequence, multidigit addition with regrouping. She decides on a criterion-referenced assessment procedure that involves preparing 25 comparable tests, each containing 10 problems that feature multidigit addition with regrouping. To maintain a moderate degree of comparability in the difficulty of the items on this "multidigit addition" test, Mrs. P. decides that all problems will present 3- or 4-digit numerals. The criterion-referenced testing procedure will involve presenting the test, along with directions, allowing 3 minutes for writing answers, and scoring performance in terms of the number of correct problems written in 3 minutes. Mrs. P. defines mastery as eight correct problems in 3 minutes on 3 consecutive days. (In a similar way, Mrs. P. would design a criterion-referenced testing procedure to assess mastery of each problem type listed in Table 1.)

Having ordered the skills embedded in the curriculum and having designed a criterion-referenced testing procedure for each skill in the instructional sequence, Mrs. P. would teach multidigit addition with regrouping and test Dolly's proficiency on this problem type on a regular basis. When Dolly achieves mastery of multidigit addition with regrouping, Mrs. P. simultanteously would shift instruction and measurement to the next teaching step: multidigit subtraction requiring regrouping. A mastery measurement graph, illustrating Mrs. P.'s measurement system for Dolly, is shown in Figure 1.

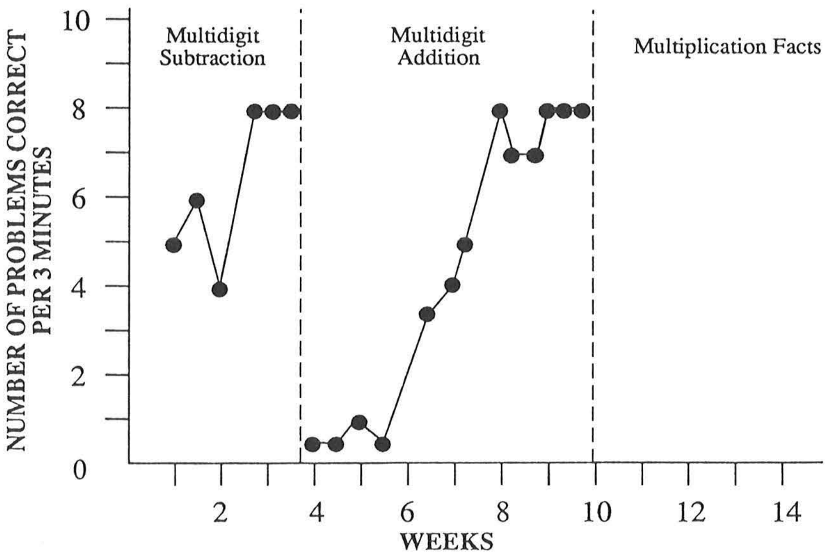

Figure 1. Example of a mastery measurement graph.

As depicted in this figure, it took 3 1/2 weeks of instructional time before Dolly demonstrated mastery of multidigit addition with regrouping. Then, when mastery of multidigit addition was achieved, Mrs. P. shifted instruction and measurement to the second step of the instructional hierarchy: multidigit subtraction. Approximately 6 weeks later, when mastery of multidigit subtraction was demonstrated, Mrs. P. began instruction on the third skill of the hierarchy, multiplication of basic facts (factors to 9). Consequently, measurement would be conducted on the criterion-referenced testing approach Mrs. P. designed to assess proficiency on multiplication facts (factors to 9).

## Curriculum-Based Measurement

As distinguished from the predominant form of curriculum-based assessment, (i.e., mastery measurement), two important characteristics of curriculum-based measurement (CBM) are (a) assessment of proficiency on skills that represent the entire, year-long curriculum and (b) reliance on standardized, prescriptive measurement methods. To clarify, let me return to the example of Mrs. P. and Dolly.

In this case, Mrs. P. maintained her goal for Dolly (i.e., proficiency on the fourth-grade computation curriculum), but she decided to rely on CBM rather than on mastery measurement. Instead of sequencing the fourth-grade computation curriculum and formulating a criterion-referenced testing procedure for each step in the instructional sequence, Mrs. P. would complete the following process.

Sheet #30
Password: JAR
Name : ______________________ Date : ______________

| | | | | |
|---|---|---|---|---|
| A<br>747<br>x 3 | B<br>5406<br>4721<br>+ 75 | C<br>1<br>x1 | D<br>2666<br>- 2647 | E<br>$9\overline{)27}$ |
| F<br>$5\overline{)800}$ | G<br>$2\frac{1}{3} + 6 =$ | H<br>0<br>x 4 | I<br>$2\frac{2}{5} - 1 =$ | J<br>5<br>x 4 |
| K<br>1670<br>4121<br>+ 86 | L<br>$7\overline{)21}$ | M<br>$4\overline{)962}$ | N<br>$7\overline{)28}$ | O<br>26.8<br>+ 13.35 |
| P<br>51<br>x 91 | Q<br>0<br>x3 | R<br>4<br>x4 | S<br>$\frac{2}{3} - \frac{1}{3} =$ | T<br>33<br>x 22 |
| U<br>$7\overline{)7}$ | V<br>16.42<br>- 3.8 | W<br>702<br>x 7 | X<br>46943<br>+ 80950 | Y<br>5<br>x 8 |

Figure 2. Example of CBM computation test.

She would list (a) the problems that constitute the fourth-grade computation curriculum and (b) the proportion of problem types that accurately represent the curriculum. For the statewide Tennessee "Basic Skills First" fourth-grade curriculum, these problem types and corresponding proportions are shown in Table. 1. This pool of problem types is the domain that Mrs. P. wants Dolly to master by June; it is Dolly's annual, year-long curriculum. Then, according to CBM methodology (Fuchs, Fuchs, & Hamlett, 1989a), Mrs. P. would use randomly generated numerals to create a series of alternate test forms. Each test would comprise 25 problems that represent the type and proportion of problems constituting the fourth-grade curriculum. One alternate form of the fourth-grade computation test is shown in Figure 2. To accomplish the test-construction process, Mrs. P. could use a computer program (Fuchs, Hamlett, & Fuchs, 1990). With this program, Mrs. P. would specify the problem types and proportions to the computer; the computer would generate the alternate forms. Then, according to standard CBM methodology (Fuchs, Fuchs, & Hamlett, 1989a), Mrs. P. would administer and score each CBM test in the following way. She would present a test and a standard set of directions to the student, and allow Dolly 3 minutes to complete as much of the test as possible. Mrs. P. would score performance in terms of the number of digits Dolly wrote correctly in 3 minutes.

Each math test samples the year-long domain in the same way; each test is an alternate form that represents the fourth-grade curriculum. As shown in Figure 2, the CBM test samples computation behaviors *across* the skills representing the fourth-grade curriculum (these skills are listed in Table 1). During the first part of the school year (i.e., in October), Dolly has poor mastery of the fourth-grade curriculum, and her scores are low on the CBM test (i.e., 18 digits correct; see scores shown in Figure 3). The total number of correct digits score on the CBM test is a performance indicator of Dolly's overall proficiency in the fourth-grade computation curriculum. The score does not communicate *which* skills in the curriculum have and have not been mastered; rather, it *indicates* that few skills are mastered. The teacher can, however, determine Dolly's specific skill profile using the CBM database. The practitioner can analyze Dolly's performance on the specific items on the CBM tests, which sample across the fourth-grade curricular skills, to determine which skills currently are mastered. When the teacher conducts such an item analysis on the CBM tests, he/she corroborates the lack of proficiency indicated by the score of 18. As shown in Table 2, which displays the profile of skills achieved at three points in time across the year, when the practitioner analyzes the responses on the items of the test, the performance indicator of 18 is associated with no mastered skills and only several partially mastered skills.

As the year progresses and instruction continues, Dolly's CBM scores increase gradually. By February, Dolly has earned scores of 45 digits correct (see Figure 3). When we analyze the responses on the CBM tests, we see that this increased score of 45 digits is associated with three mastered skills, five partially mastered, and only two nonmastered skills in the fourth-grade curriculum. Then, as time passes and additional instruction occurs, Dolly gains proficiency on the fourth-grade curriculum; her performance indicator continues to increase to 55 by April (see Figure 3), and the profile of fourth-grade skills mastered concurrently improves (see Table 2).

Within CBM, the performance indicators are presented in graphic form. For example, the graph in Figure 3 shows Dolly's scores on the CBM tests across time. As the year progresses, Dolly's scores increase. The slope of Dolly's scores across time represents Dolly's overall learning rate in the fourth-grade curriculum. As the performance indicator (or CBM score) increases, Mrs. P. knows that Dolly's overall proficiency in the fourth-grade curriculum has increased, and she has confidence that Dolly's mastery of specific fourth-grade skills also is improving.

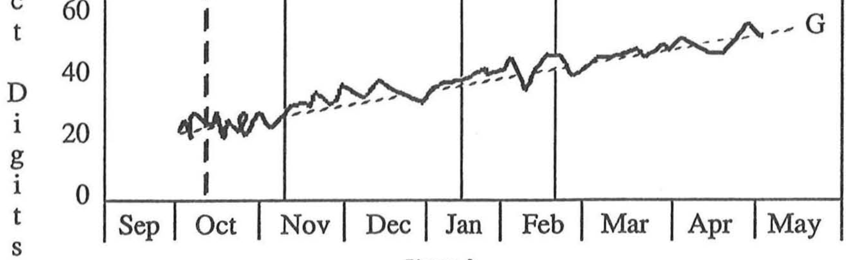


Figure 3.

**Table 2**
Skills Profile at Three Points in Time

| Date | Mastered | Partially Mastered | Nonmastered | Not Attempted |
|---|---|---|---|---|
| October | | Multidigit addition, regrouping<br>Multiplication facts<br>Multiplication 1- or 2-digit, regrouping | Multidigit subtraction, regrouping<br>Multiplication, no regrouping<br>Division facts<br>Dividing 3- by 1-digit, no remainder | Dividing 2- or 3- by 1-digit, remainder<br>Adding/subtracting simple/mixed fractions, no regrouping<br>Adding/subtracting mixed decimals to hundredths |
| February | Multidigit addition, regrouping<br>Multiplication facts<br>Multidigit subtraction, regrouping | Multiplication 1- or 2- digit, regrouping<br>Division facts<br>Dividing 3- by 1-digit, no remainder<br>Dividing 2- or 3- by 1-digit, remainder<br>Adding/subtracting mixed decimals to hundredths | Multiplication, no regrouping<br>Adding/subtracting simple/mixed fractions, no regrouping | |
| April | Multidigit addition, regrouping<br>Multiplication facts<br>Multidigit subtraction, regrouping<br>Division facts<br>Dividing 2- or 3-digit by 1-digit, remainder<br>Adding/subtracting mixed decimals to hundredths | Multiplication, no regrouping<br>Multiplication 1- or 2-digit, regrouping<br>Dividing 3- by 1-digit, no remainder | Adding/subtracting simple/mixed fractions, no regrouping | |

### Important Distinctions Between Mastery Measurement and CBM

Five important distinctions exist between mastery measurement and CBM. These salient differences are (a) the scope of skills upon which measurement is focused, (b) the extent to which generalization and maintenance are assessed, (c) the degree of constancy in measurement across time, (d) the reliance of the measurement on instructional hierarchies, and (e) the methods by which measurement methods are developed. An explanation of each of these differences follows.

*Scope of skills for measurement.* Mastery measurement and CBM are essentially different because of the scope of skills encompassed within these two forms of measurement. Specifically, mastery measurement is relatively narrow; it focuses measurement on single skills (or small clusters of skills) at a time. By contrast, CBM is relatively broad; it focuses measurement on a large domain of skills, representing the curriculum to be mastered over the course of a school year.

Mastery measurement focuses instruction and measurement on a series of short-term instructional objectives; therefore, instruction and measurement are linked together. An advantage of this linking is that the assessment data should be highly sensitive, or responsive, to instructional effects. This indicates strong instructional validity (Yalow & Popham, 1983). Nevertheless, a potential disadvantage of a close connection between measurement and instruction is that the measurement framework is restricted. Scores may reflect the student's skill in computing only in the narrow framework within which testing occurs (i.e., when all problems require use of the same multidigit-regrouping addition algorithm). So, the content validity, reflecting the extent to which the measurement mirrors the domain—computing problems in natural or mixed presentation—may be reduced. Also, the relation between progress through an instructional sequence and socially important outcomes, such as standardized, commercial achievement test performance, is uncertain.

In contrast, CBM focuses on the long-term goal. That is, rather than measuring student mastery on a series of changing instructional objectives, CBM focuses measurement on the relatively broad, annual curriculum. The disadvantage associated with such a broad focus is the loss of potential instructional validity. Compared to mastery measurement, where the teacher tests performance on the immediate instructional objective, CBM samples content across the year-long curriculum. Consequently, CBM may be less sensitive than mastery measurement to student change as a result of current instruction (Fuchs

& Deno, in press). However, compared to traditional measurement, where performance samples behavior across both grade levels and curricula at one moment in time, CBM provides information that (a) is sensitive to instructional effects (Marston, Fuchs, & Deno, 1985) and (b) can be used to improve instructional decision making (Fuchs & Fuchs, 1990).

Also, as can be anticipated in light of the foregoing discussion, CBM's focus on long-term goal measurement offers certain advantages over mastery measurement. Because CBM describes student performance in terms of proficiency on the annual curriculum, both its content and criterion validity are stronger than mastery measurement (Fuchs & Fuchs, 1986).

*Retention and generalization of skills.* A second key distinction between mastery measurement and CBM is the extent to which the measurement assesses retention and generalization of skills. With mastery measurement's close connection between testing and instruction, mastery measurement does not automatically assess retention and generalization of skills. When Dolly demonstrates mastery of multidigit addition with regrouping (and when measurement and instruction simultaneously shift to subtraction with regrouping), we have no automatic index of the extent to which Dolly retains mastery of multidigit addition. Conversely, while Mrs. P. focuses instruction and testing on multidigit addition, we have no indication of the extent to which Dolly may generalize her increasing skill in multidigit addition to other dimensions of the curriculum. For example, as Dolly gains mastery of multidigit addition with whole numbers, she may acquire skill in mixed addition of decimals to the hundredths place. Yet, a mastery measurement system will not index this generalization. As this illustrates and as Goodstein (1982) has described, closely linking the instructional format to assessment (or narrowly defining the content-x-format domain of criterion-referenced/mastery measurement) may create problems, including the failure to index retention and generalization learning events.

In contrast to mastery measurement, CBM offers the advantage of automatically assessing retention and generalization of skills. As Dolly improves her skill in multidigit addition with regrouping, the CBM performance indicator should increase, because Dolly's increased proficiency allows her to compute the multidigit addition problems with regrouping (and therefore more digits) correctly on the CBM tests. However, if Dolly fails to retain mastery of multidigit addition with regrouping when multidigit subtraction with regrouping instruction begins, Dolly's CBM score should decrease. This would occur because

Dolly no longer would compute the multidigit addition with regrouping problems on the CBM tests correctly. Therefore, CBM is sensitive to retention because it samples skills *across* the annual curriculum.

Conversely, if Dolly generalizes learning to new skills when multidigit addition with regrouping instruction occurs, Dolly's performance indicators should increase, because opportunities for computing untaught problem types are provided on the CBM tests. In this way, CBM indexes generalization. This sensitivity of measurement to retention and generalization learning may be critical when CBM is used to monitor the development of basic skills for handicapped populations. These low-achieving pupils frequently have poorly developed strategies for maintaining and transferring skills (Anderson-Inman, Walker, & Purcell, 1984; White, 1984).

*Constancy in measurement across time.* A third difference between mastery measurement and CBM is the extent of constancy in measurement across time. Mastery measurement requires a shift in measurement each time a skill is mastered; CBM maintains a constant measurement focus across the year.

As shown in Figure 1, with the regular shifts in mastery measurement across time, we can determine an acquisition rate for multidigit addition with regrouping and we can estimate a separate learning curve for acquisition of multidigit subtraction with regrouping. However, it is impossible to summarize an overall learning rate *across* the different skills in the curriculum. This is because different skills, measured at different times during the school year, are not of equal difficulty and do not represent equal curriculum units. For example, research indicates that acquisition of subtraction skills is more difficult than mastery of addition skills. Consequently, one would not expect different skills (even seemingly analogous skills such as multidigit addition with regrouping and multidigit subtraction with regrouping) to be acquired in equivalent times. These unequal curriculum units, along with the shifts in measurement and the resultingly limited summaries of learning rate, appear to reduce the usefulness of mastery measurement.

With CBM, teachers may monitor students' basic skills development across a school year without any shifts in measurement. Because CBM tests sample across the entire year-long curriculum, test difficulty remains constant across the school year. As shown in Figure 3, the difficulty of the CBM tests Dolly took in November is comparable to the difficulty of the tests she took in March. It is Dolly's *proficiency*, not the test difficulty, that increases. However, with mastery measurement, the measurement domains and the difficulty of testing material continually change as the instructional content changes. CBM avoids these shifts in measurement domains, and this constancy associated with CBM permits summaries of student learning rates across time. The CBM database can be used to compare the effectiveness of different instructional components introduced at different times during the year (see subsequent discussion).

*Reliance on instructional hierarchies.* Another key distinction between mastery measurement and CBM is the extent to which they rely on instructional hierarchies to determine measurement. In order to establish mastery measurement systems, teachers are required to specify instructional hierarchies that dictate the sequence for instruction and measurement. Most instructional hierarchies rely on "scope and sequence" charts (see Salvia & Hughes, 1990, for procedures for specifying instructional hierarchies within mastery measurement). Such charts tend to be long and detailed, and require teachers to group across skills (Salvia & Hughes, 1990). Additionally, scope and sequence charts typically are based on logical, rather than empirical, analyses of skills development. The appropriateness of logically determined sequences of instruction for students, especially handicapped pupils who do not progress along predictable developmental sequences, is unknown. Moreover, as demonstrated in the discussion that follows, when instructional hierarchies *determine* measurement, teachers cannot use assessment information to evaluate the effectiveness of alternative instructional approaches.

As opposed to mastery measurement, CBM does not require teachers to specify instructional hierarchies before measurement occurs. To set up a CBM system, a teacher identifies the annual domain on which he/she expects the student to be proficient by June. This offers certain advantages. First, the difficult task of compartmentalizing and ordering the curriculum is circumvented. This eliminates teacher effort, and avoids possible errors in specifying instructional chunks and sequences that eventually may prove troublesome to individual student growth.

Second, in sharp contrast to mastery measurement, CBM does not *determine* instruction. The structure of mastery measurement specifies the order in which instruction must proceed, and one cannot progress to subsequent skills until mastery of the current skills is demonstrated. Moreover, as illustrated in the work of Salvia and Hughes (1990), the mastery measurement framework also typically results in a skills-oriented approach to instruction, and the order in which skills are taught is determined by measurement. With mastery measurement, the independent variable (instruction) and the dependent variable (measurement) are tied together, with both simultaneously focused on skills. With CBM, measurement (the dependent variable) is not tied to and determined by the current instructional focus or procedure (the independent variable); therefore, measurement and instruction are not confounded. Because of this, CBM offers the advantage of permitting teachers to experiment with contrasting instructional chunks, sequences, and/or procedures: Teachers use the CBM database as the dependent variable by which they evaluate the effectiveness of contrasting instructional strategies.

*Development of tests.* The fifth feature that differentiates mastery measurement and CBM is test development procedures. Mastery measurement relies primarily on the use of teacher-made criterion-referenced tests. Such teacher-made criterion-referenced tests have unknown technical characteristics. And the time-consuming and costly nature of reliability and validity studies makes it difficult, if not impossible, to investigate the psychometric characteristics of teacher-constructed measures. Additionally, even when teachers rely on commercial criterion-referenced tests for mastery measurement, psychometric characteristics are uncertain. Hambleton and Eignor (cited in Berk, 1982) evaluated 11 popular, commercially available criterion-referenced tests. They found that these tests could be characterized as follows:

- About half of the publishers included information about the qualifications of individuals who prepared the objectives on which the tests were based.
- Item representativeness could not be established because of the absence of domain specifications.
- For item analysis, there were two problems: Too little explanation was offered for the choice of particular item statistics and for the specifics of item statistics usage; and item statistics were used in test construction, thereby "biasing" the content validity of the test in unknown ways.
- Test score reliability was not handled well in most manuals.
- Inappropriate, or no, information relative to the stated uses of the test scores was offered.
- Rationales and procedures for setting cutoff scores were not offered, and evidence usually was not provided for the validity of cutoff scores (e.g., did examinees classified as masters typically perform better than those classified as nonmasters on some appropriate external criterion measure?).
- Factors affecting the validity of scores were not offered in any manuals.

-Few manuals introduced the notion of error in test scores or classifications of examinees to mastery states.

These findings, based on examination of criterion-referenced test manuals, are corroborated by empirical work. Tindal, Fuchs, Fuchs, Shinn, Deno, and Germann (1985) conducted reliability and validity studies on criterion-referenced tests associated with four popular basal reading series. Findings indicated variable reliability and validity coefficients, with many indices failing to reach acceptable levels. Consequently, commercial criterion-referenced tests frequently fail to provide information with documented reliability and validity.

In contrast to typical mastery measurement approaches, a comprehensive research program (Deno & Fuchs, 1987; Shinn, 1989) has investigated the psychometric characteristics of alternative methods for sampling test stimuli from curriculum, administering and scoring tests, and summarizing and evaluating scores in prescriptive ways. From this research, a standard CBM methodology has been formulated (Mirkin et al., 1984). Consequently, when teachers have determined the curriculum they expect students to master over the course of the school year, CBM prescribes methods for creating, administering, scoring, and using tests that result in reliable and valid descriptions of students' basic skills growth in reading, spelling, written expression, and computation. This standardized, prescriptive measurement within CBM, with documented reliability and validity, contrasts sharply with the unknown psychometric features of the teacher-made criterion-referenced tests used within mastery measurement.

## USING CURRICULUM-BASED MEASUREMENT TO DEVELOP EFFECTIVE INSTRUCTIONAL PROGRAMS

Research supports three strategies for using curriculum-based measurement (CBM) to assist teachers in developing instructional programs. First, teachers can use CBM to monitor the appropriateness of the goals they set and to ensure the use of realistic, but ambitious, goals. Second, CBM can be used to determine the adequacy of student progress, to determine whether instructional programs require adjustment, and to compare the effectiveness of alternative programmatic components. Finally, CBM databases can be used to draw profiles of skill strengths and weaknesses, in order to assist teachers in determining the nature of effective programmatic modifications. In the following sections, each of these applications is described and the relevant research base is reviewed.

## Using CBM to Monitor and Adjust Goals

Research substantiates the effectiveness of using goals to improve instructional outcomes. Summarizing across a variety of goal-writing procedures and research methods, Hartley and Davies (1976) found that teaching with goals enhances student achievement. McNeil (1967), for example, demonstrated that teachers who employed behavioral objectives produced better academic growth with their students and were judged to be more successful in applying learning principles, compared to a control group of teachers who did not use goals.

The relevant literature suggests that one way in which goals may mediate enhanced achievement outcomes is by structuring evaluation activities. A well-written goal defines the parameters of measurement: The goal specifies the anticipated observable performance that is desired, the conditions under which the behavior will be demonstrated, and the criteria against which to judge performance (Bloom, Hastings, & Madaus, 1971; Gagne, 1964; Mager, 1975). Adding this structure to the evaluation process may help teachers generate frequent, relevant student performance data. With ongoing feedback to practitioners and students, teachers can formulate more effective instructional programs (Jenkins, Deno, & Mirkin, 1979), and students can recognize their own successful learning strategies more readily (Bandura, 1982; Peckham & Roe, 1977; Rosswork, 1977).

CBM attempts to take advantage of potential benefits associated with the use of goals. Within CBM, the structure of the goal establishes key dimensions of the measurement/evaluation system. First, as the teacher selects the goal, she specifies the point within the curriculum where the student is expected to be proficient by year's end. This level becomes the measurement pool from which stimuli for testing are drawn. Second, when setting the goal, the teacher simultaneously indicates the performance criterion she is equating with "proficiency." This performance criterion creates the structure against which the adequacy of student progress is judged within CBM.

Let us say, for example, that Mrs. P. determines she wants a second student, Michael, to be proficient in Grade 3 of the computation curriculum by the end of the school year. Using CBM, Mrs. P. would measure Michael's performance on an alternate test, comprising 25 problems that represent the type and proportion of problems in the same way each time she tested Dolly's proficiency in the curriculum. Let us also say that Mrs. P. equates "proficiency" for Michael in this curriculum with a score of 20 digits correct by April 15. Using CBM, Mrs. P. would set up a monitoring graph to create a record of Michael's progress and to evaluate the adequacy of Michael's growth. As shown in Figure 4 (top panel), this graph displays Michael's initial, or baseline, performance in the target Grade 3 curriculum (see dots that show scores of 5, 9, and 6); it shows the goal (see the "G" placed at the desired score

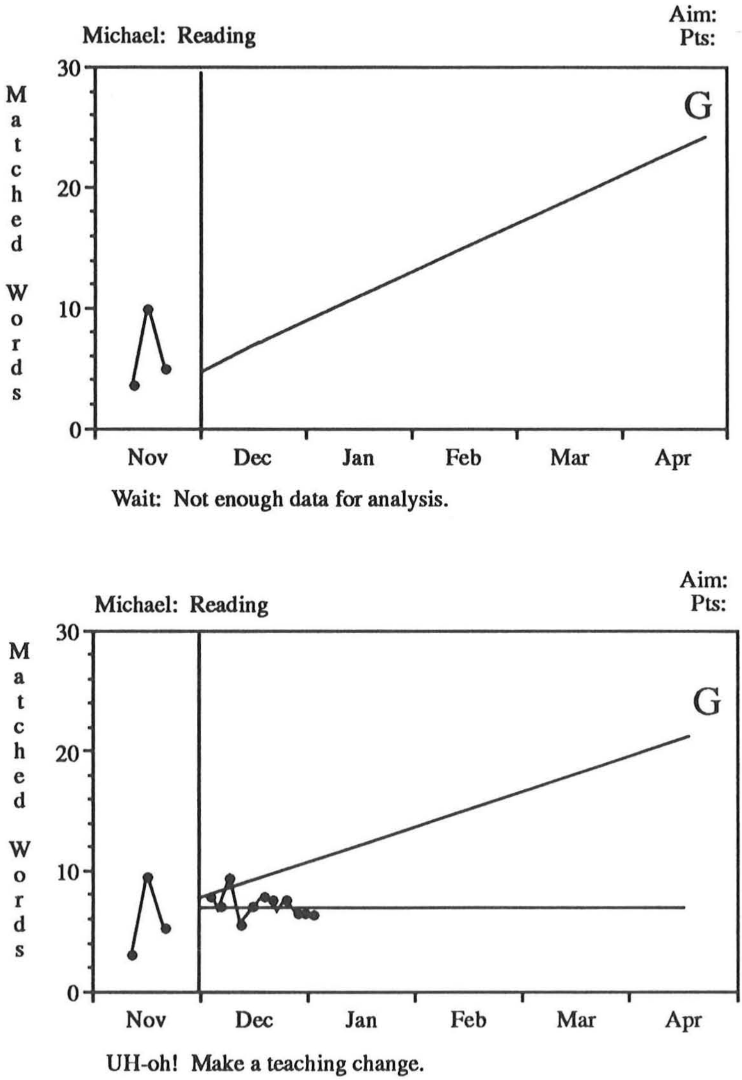


Figure 4. Example of CMB graphs. The top panel shows baseline, goal, and goal line; the bottom panel illustrates data for which the appropriate decision is for the teacher to make a teaching change.

of 20 on April 15); and it illustrates a "moving goal" (see the broken diagonal line) that indicates (a) the rate at which Michael will have to improve in order to attain the goal and (b) the target score on any given date.

Within typical CBM practice, the goal structures the evaluation process in the following way. When the student's actual rate of progress falls below the rate necessary for goal attainment, the rate of the student's progress and the effectiveness of the student's program are judged inadequate. In this case, CBM decision rules dictate that a teaching change is required. Figure 4 (bottom panel) shows an example of such a decision. Here the student's actual rate of progress, indicated by the solid diagonal line, is less steep than the desired rate of progress for goal attainment, indicated by the broken diagonal line. As illustrated, the decision in this case would be for the teacher to modify the instructional program in order to stimulate student progress.

As this discussion should make clear, the performance criterion specified in the goal becomes critical in the instructional decision-making process. Within the context of programming for handicapped or other low-achieving students, where the need for quality instructional programming is essential, the most critical potential problem associated with the performance-criterion-setting process may be the following: When teachers set goals that are unambitiously low, few if any recommendations for instructional improvements will be made.

Moreover, research indicates that unambitious goal setting within CBM relates to relatively poor student achievement. Fuchs, Fuchs, and Deno (1985) conducted a post-hoc analysis of a database in which each teacher, along with their four mildly to moderately handicapped pupils, had been assigned randomly to either a CBM or a control group condition for a 4-month study in the area of reading (Fuchs, Deno, & Mirkin, 1984). In this post-hoc study, student graphs were inspected after the completion of the CBM implementation. On the basis of inspecting graphs and looking at teachers' setting of goals and students' final performance levels, the 58 students in the CBM group were divided into three goal ambitiousness conditions: a highly ambitious goal group, a moderately ambitious goal group, and a low ambitious goal group. Students also were divided into two goal mastery conditions: those who had mastered and those who had not mastered their goals.

Three types of achievement outcomes were studied: (a) the Passage Reading Test, a measure that requires reading behavior similar to that required in the CBM tests; (b) the Stanford Diagnostic Reading Test, Structural Analysis subtest, a measure of decoding skills; and (c) the Stanford Diagnostic Reading Test, Reading Comprehension subtest. A multivariate analysis of covariance, with appropriate follow-up analyses, indicated the following. The ambitiousness with which the goals were established was associated positively with student achievement. On two achievement measures, with pretreatment achievement levels statistically controlled, students for whom teachers set highly and moderately ambitious goals achieved better than students whose goals reflected relatively unambitious goals. On a third achievement measure, students with highly ambitious goals performed better than students for whom moderately ambitious and low goals were set. Furthermore, there were no effects associated with goal mastery. That is, students who met their goals and students who did not meet their goals achieved in comparable fashion. It was the level of goal *ambitiousness*, not goal *attainment*, that was associated with student achievement.

Based on these results, it appears that the selection of an appropriately ambitious, but realistic, performance criterion appears to be critical within CBM instructional decision making. Despite this importance, few satisfactory strategies for identifying appropriate performance criteria have been formulated. One potential solution to the goal-setting problem with CBM, referred to a dynamic goal setting, has been explored recently.

During the 1986-1987 academic year, Fuchs, Fuchs, and Hamlett (1989a) conducted a study designed to test the effectiveness of an innovative CBM goal-setting strategy, "dynamic" goal setting. In this study, participants were 30 special education teachers who taught self-contained and resource programs for students in Grades 2—9. Teachers selected two mildly handicapped students with IEP math goals. Then, teachers were assigned randomly to three treatment groups: dynamic goal CBM, static goal CBM, and control. The control teachers monitored student progress using conventional special education practice, including unit tests, correction of assignments, and unsystematic observation of student performance. The teachers in both CBM groups did the following. For 15 weeks, each teacher employed CBM to track their two pupils' progress toward math goals. The CBM system was rooted in the Tennessee Basic Skills First Math Program (BSF). The math computation objectives tested at each grade level within the BSF were listed. Teachers inspected these lists and determined an appropriate grade level on which to establish each student's goal. This level included the pool of math objectives the teacher hoped the student would master by year's end.

Using a standard measurement task, teachers were required to assess each pupil's math performance at least twice weekly, for 2 minutes, each time on a different probe representing the type and proportion of problems from the BSF goal level they had selected. That is, if the teacher chose the third-grade level of the curriculum, the teacher was provided with 50 alternate test forms, each of which sampled the BSF third-grade computation objectives in the proportion tested on the BSF third-grade criterion-referenced end-of-year test. Each test could be conceptualized as a short form of the BSF third-grade computation test. Consequently, as teachers monitored pupil progress on these tests, they could estimate progress toward mastery of the corresponding level of the BSF end-of-year tests.

Each test was scored in terms of the number of correct digits written in 2 minutes. For half the students in each CBM group, scores were automatically collected using computers and saved to disk; for the other half, scores were collected by teachers and entered into a data-management software program by teachers. However, all testing procedures were completely analogous, and no outcome differences were associated with this administration factor (Fuchs, Hamlett, & Fuchs, 1987). Once each week, teachers used data-management software to review their students' assessment profiles. The software automatically graphed the scores, drew a goal, a goal line, and a regression line of best fit depiciting the student's actual slope of improvement. Additionally, the software applied a set of decision rules. If the regression line was less steep than the goal line, the decision provided to the teacher read, "Uh-oh! Make a teaching change." When the regression line was steeper than the goal line, one of two possible decisions came up, depending on the teacher's experimental condition.

Within the *static goal* CBM group, when the student's actual rate of improvement exceeded the rate anticipated in the goal line, the decision read "OK! Collect more data." The data pattern suggested that the student's rate of progress was acceptable with respect to goal attainment, and that the corresponding instructional program looked effective. Thus, the message indicated that the teacher should keep the current instructional program intact and continue data collection. The teachers always were free to increase their goal, but they never were directed to do so. Figure 5 (top panel) shows a graph depicting satisfactory progress, and the message that would have been delivered within the static goal CBM condition.

Within the *dynamic goal* CBM group, when the student's actual rate of improvement exceeded the rate anticipated in the goal line, the decision read "OK! Raise the goal to X" (where X = the student's predicted performance at the end of the study, based on the student's current rate of progress). Again, the data pattern suggested that the

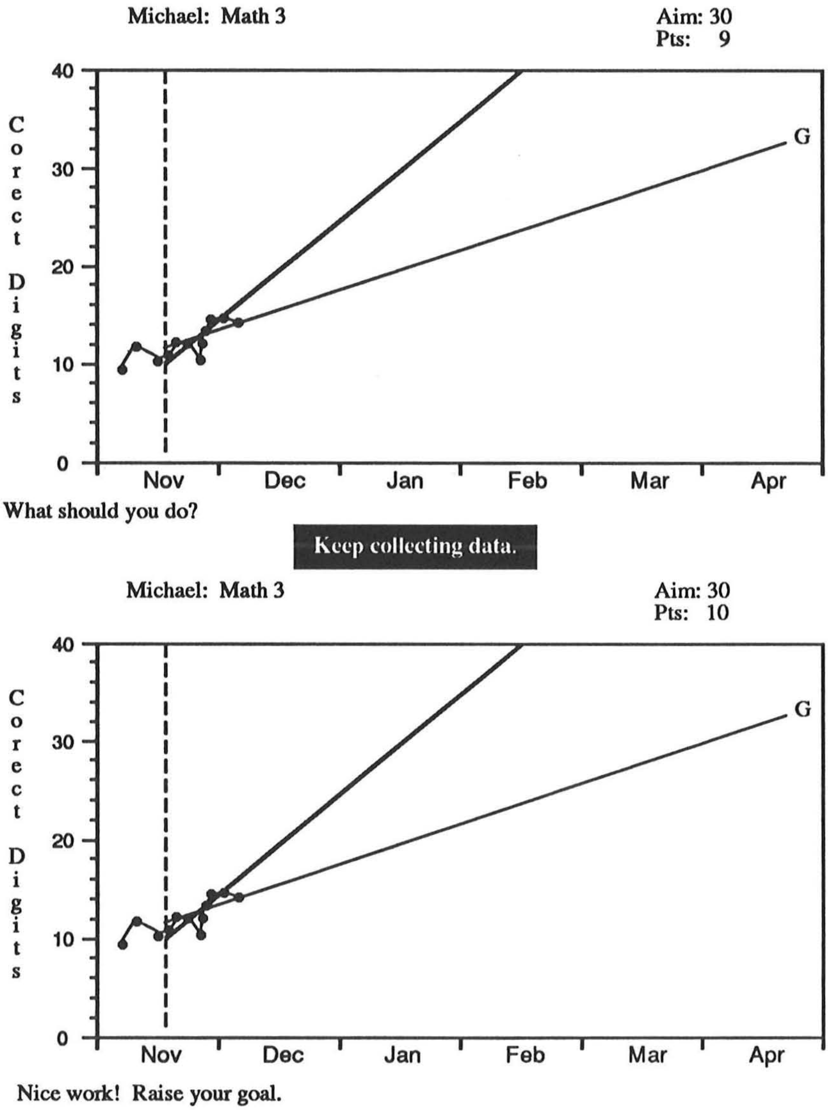


Figure 5. Example of CMB graphs, where actual performance exceeds the progress anticipated by the teacher during goal setting. In the top panel, a static goal decision is shown, which suggests that the teacher continue to collect data. In the bottom panel, a dynamic goal decision is shown, which suggests that the teacher raise the goal.

student's rate of progress was acceptable with respect to goal attainment, and that the corresponding instructional program looked effective. The message indicated that the teacher should maintain the current instructional program and continue data collection. However, the teacher also was required to raise the goal. Figure 5 (bottom panel) shows a sample graph, illustrating satisfactory progress, with the message that corresponded to the dynamic goal CBM condition. By raising the goal, the teacher accomplished two things. First, she always adjusted the goal to correspond to the student's actual rate of progress or better; the goal was not allowed to reflect a progress rate lower than that which the student could achieve. Second, and perhaps more important, by adjusting the goal upward, the teacher was simultaneously establishing a more ambitious criterion for subsequent decisions concerning the adequacy of student progress and the instructional program. With a raise in the goal, the likelihood increased that the teacher would receive a recommendation for a teaching change in subsequent evaluations.

Two types of outcomes associated with this study are especially interesting. One type of outcome concerns teachers' use of goals; the other, student achievement. With respect to use of goals, teachers in the dynamic goal CBM group made more goal increases than teachers in the static goal CBM group. Given the dimensions of the different CBM conditions, this finding is not surprising. What is more interesting is the magnitude of effect. Within the dynamic goal group, teachers made an average of .60 goal increases; that is, they increased goals for more than one out of every two pupils. In the static goal group, only one teacher, for one of her pupils, spontaneously increased a goal in response to the student's data.

This finding is important for several reasons. First, it suggests that, despite the potential importance of ambitious goals, special educators' typical goal-setting standards may underestimate many students' potential. The study procedures allowed teachers to establish their initial goals freely, in line with the progress rates they deemed ambitious but realistic. However, with these initial goals, teachers in the dynamic goal group were required to increase goals for more than one out of every two pupils. This goal-increasing behavior was prompted by students exceeding the rates of progress teachers had anticipated. This goal-increasing rate, in response to students exceeding teachers' initial expectations, has been corroborated in additional studies we have conducted, in other academic areas. During the 1987--1988 school year, we used the dynamic goal condition in reading, spelling, and math. In these three academic areas, respectively, teachers were required to increase goals for 4 out of every 10 pupils, 6.5 out of every 10 pupils, and 4 out of every 10 pupils. It appears that teachers may systematically underestimate handicapped students' potential to grow.

In addition to demonstrating that teachers' goals may underestimate potential progress rates, these findings indicate that without systematic prompting to raise goals, practitioners cannot be expected to do so. For example, among the 20 students participating in the static goal group, there was only one instance of a teacher raising a goal. Therefore, similar to research that indicates the importance of decision rules to prompt teachers to make instructional changes, it appears that decision rules prompting teachers to raise goals may be necessary.

The second major outcome of interest in the Fuchs, Fuchs, and Hamlett (1989a) study concerns student achievement. Concurrent with teachers' goal-raising behavior was differential student achievement. Students in the dynamic goal CBM group achieved better than the controls during posttesting on a standardized computation achievement test (with pretest performance controlled statistically). However, the achievement of the static goal CBM group did not exceed that of the controls. The effect size associated with the dynamic goal CBM procedures was .52, or approximately one-half standard deviation. This indicates that, in terms of the standard normal curve and an achievement test scale with a population mean of 100 and standard deviation of 15, one might expect the use of CBM with dynamic goals to increase the typical achievement outcome score from 100 to approximately 107.5. This finding supports previous research in psychology indicating that adults in work settings perform better with difficult goals. Additionally, findings corroborate a post-hoc special education analysis (Fuchs et al., 1985) where teachers who employed more difficult CBM goals effected better student achievement.

The Fuchs et at. (1989a) study, therefore, contributes to the CBM literature by providing an example of a workable methodology the special education community might employ for empirically deriving ambitious, but realistic, goals. A persistent problem for special education has been that during the IEP development process, before the efficacy of special education intervention has been established for a particular student, it is difficult, if not impossible, to anticipate the scope of attainable, but ambitious, goals. The Fuchs et al. study provides a process by which goals can be developed dynamically, so that progress toward mastery is monitored closely and goals are adjusted upward whenever possible. Given the finding that such goal adjustment, specifically, and goal ambitiousness, generally, may enhance student achievement, the special education community might consider adoption of CBM systems that incorporate dynamic goal-setting procedures.

### Using CBM to Judge the Adequacy of Student Progress and to Adjust Instructional Programs

Using CBM to monitor the appropriateness of instructional goals and to adjust goals upward whenever possible represents one means by which CBM can be used to assist teachers in their instructional program development. A second key way in which the CBM database can be used to enhance instructional programs is to provide the essential information with which teachers can determine (a) the adequacy of student progress, (b) the effectiveness of the current instructional program, and (c) the relative efficacy of alternative programmatic components.

Each CBM score is a performance indicator, representing the student's overall proficiency in curriculum on which measurement is conducted. Increasing scores indicate enhanced proficiency; decelerating or flat scores signify a lack of growth. As discussed previously in this chapter, when a teacher sets a goal and thereby establishes a moving goal line for a particular student, he/she simultaneously sets a minimally acceptable rate of improvement for the student, as indexed by the performance indicators. Consequently, when a student's actual rate of growth (see solid diagonal line in Figure 4) is flatter than the student's anticipated rate of growth (see broken diagonal goal line in Figure 4), a student's growth rate and the student's instructional program are judged inadequate. At this point, a recommendation is provided to make a teaching change, in order to stimulate better growth.

A series of studies indicates the importance of this "instrumental" use of the CBM database to assist teachers in judging the adequacy of student progress in order to develop enhanced instructional programs as necessary. For example, in a meta-analysis of systematic formative evaluation studies, Fuchs and Fuchs (1986) found that the use of decision rules to stimulate teachers' use of monitoring databases for programmatic development resulted in better student achievement. Fuchs et al. (1988) found a relation between student achievement and teachers' compliance with decision rules requiring teaching changes when student rates of progress were inadequate.

Additionally, in a post-hoc analysis of teachers' use of CBM in reading, Fuchs, Fuchs, and Hamlett (1989b) identified differential patterns of student achievement associated with teachers' instrumental use of CBM databases in order to formatively develop better instructional programs. During the 1986--1987 school year, 29 teachers were assigned randomly to two treatment groups: a control group and a group that used CBM to monitor their students' reading growth. In the control group, 17 mildly handicapped students participated; in the CBM group, subjects were 36 students with mildly handicapping conditions.

In the control group, teachers used conventional special education practice to monitor student growth. As indicated on a posttreatment questionnaire, this conventional practice included unsystematic observation of student performance during lessons and grading of worksheets and other assignments.

The CBM teachers monitored student progress using CBM. Specifically, they identified curriculum levels in which student progress would be monitored and set a performance criterion for acceptable performance at the end of the 15-week study. Twice each week, teachers measured student performance with CBM. One half of the CBM teachers used a standard recall measure to monitor student growth; the other half, a standard cloze task. Additionally, within each type of measurement group, one half of the teachers measured student performance by hand and entered student scores into a data-management program; for the other half, student measurements were collected and scored automatically by computers and scores were saved directly for the data-management disk. Preliminary analyses indicated no effects associated with the type of measure condition or the type of administration factor.

Each week, teachers employed data-management software (Fuchs et al., 1987) that automatically stored and graphed the student scores, applied a set of CBM decision rules to the graphed database, and communicated decisions to teachers based on the CBM decision rules. As in the Fuchs et al. (1989a) study, the decision rules were as follows: If the student's actual rate of improvement was less steep than the goal line, the decision was to initiate an instructional change; if the student's actual rate of progress was steeper than the goal line, the decision was to increase the goal.

Following the completion of the 15-week study, the graph of each CBM student was inspected to create two CBM implementation groups: the measurement-alone group and the measurement-with-evaluation group. For the purpose of creating these two CBM subgroups, measurement was defined as administering, scoring, and graphing the curriculum-based measures on a routine basis. Evaluation was defined as the teacher introducing at least one instructional modification in response to the database and maintaining that modification for at least 2.5 weeks. Maintenance of the modification was included as a criterion to insure that an instituted modification was in effect long enough to influence student performance.

Students were placed in the measurement-alone CBM group when their graphs showed that, although CBM measurement had occurred, the CBM database had not been used to evaluate the effectiveness of instruction and no instructional changes had been introduced in order to enhance student learning. For these students, only one viable, unchanging instructional phase had been implemented over the 15-week study. In this measurement-alone group, there were 15 students, involving nine teachers.

The remaining 21 students were placed in the measurement-with-evaluation CBM group. These students' graphs showed both that CBM data had been collected and that teachers had used the databases to evaluate and enhance instructional effectiveness. Among these students, six had three viable, different instructional phases, each implemented for at least 2.5 weeks, and 15 had two viable, different instructional phases, each implemented for at least 2.5 weeks.

Figure 6 shows two sample graphs. In the top panel, the vertical lines on the graph indicate that the teacher responded the CBM database to determine the adequacy of student growth and to develop better instructional programs; this graph would have been placed in the measurement-with-evaluation group. The bottom panel shows similar data, but the graphs lack vertical lines (i.e., no instructional changes were instituted in response to the database). Yet, as can be seen, the data pattern indicates that the teacher should have (but failed to) responded to the data instrumentally to introduce instructional changes. This graph would have been placed in the measurement-only group.

Two types of measures were used to compare the achievement of the two CBM implementation and the control groups. The first measure was a well-accepted, broadly used outcome, the Stanford Achievement Test's Reading Comprehension subtest, which was administered on a posttreatment basis and for which scores were statistically controlled using a recall measure that had been administered prior to the study. The second measure was the slope of the actual CBM database, or the rate of weekly increase in the CBM scores collected by the teachers or computers.

Results corroborated the importance of the evaluation component of CBM for effective instructional programming. Although teachers in both implementation groups set up their measurement systems and actually measured student performance using CBM comparably well, as indexed on the fidelity of treatment measure, important differences

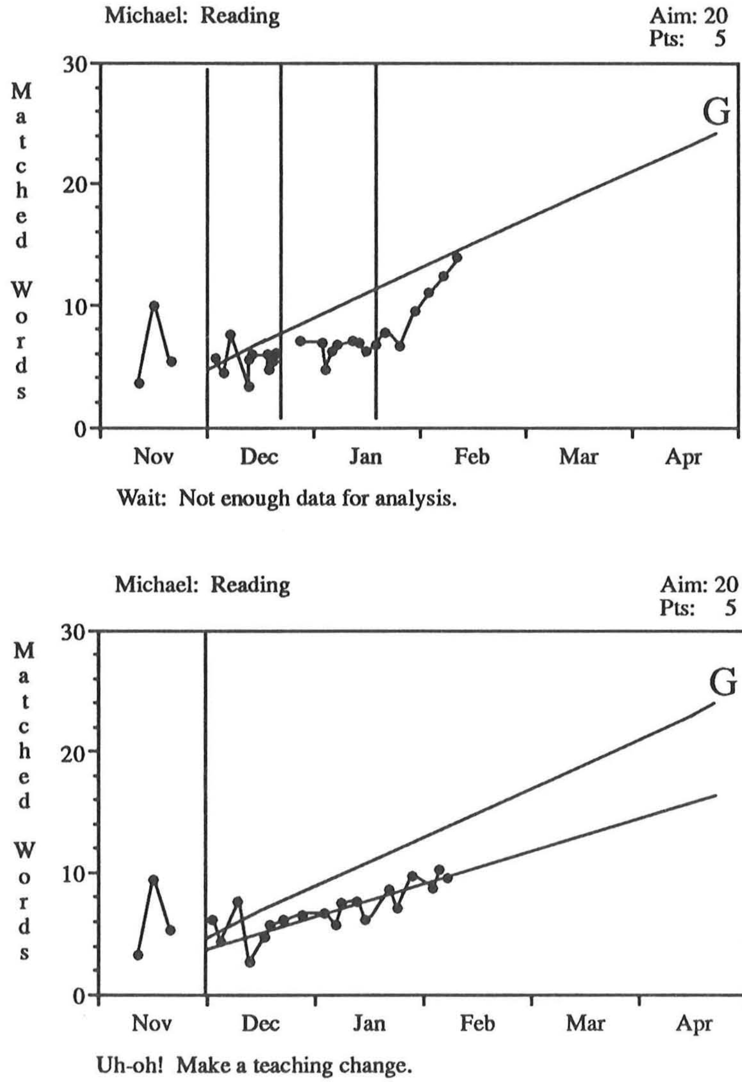


Figure 6. Example of CBM graphs. Top panel indicates that the teacher has used the databases to formulate instructional decision, as indicated by the vertical intervention lines. The bottom panel shows similar data; however, the teacher has not used the database to determine when to introduce teaching changes in order to effect greater student growth.

were associated with the CBM implementation groups.

In terms of the global, widely accepted reading comprehension measure (the Stanford Achievement Test), findings indicated that, when teachers implemented both the measurement and evaluation components of CBM, their students achieved better (in terms of regressed adjusted scores) than the control group students. However, when teachers implemented only the measurement component of CBM, without using the database to determine when instructional improvements were warranted, student achievement did not reliably exceed that of the control group. Further, the effect size for the measurement-with-evaluation CBM group was twice as large as that of the measurement-only group.

Additionally, although the difference between the measurement-only and the measurement-with-evaluation CBM groups was not reliably different on the global Stanford Achievement Test, differences on the more direct CBM index indicated that the measurement-with-evaluation group's achievement did exceed that of the measurement-only group. The effect size was .86.

Consequently, findings support the importance of the evaluation component of CBM. With the CBM evaluation component, teachers can determine when student rates of progress are less than adequate and when program changes are warranted. When teachers not only collect CBM data, but also use CBM indicators of student growth to evaluate the effectiveness of instructional programs and to experiment with alternative instructional components, student achievement appears to be enhanced.

### Using CBM to Determine the Nature of Effective Instructional Modifications

As discussed, the first strategy for using CBM databases in order to enhance teachers' instructional planning involves relying on the graphed performance indicators to monitor the appropriateness of the student's goal and to adjust the goal upward whenever necessary to ensure appropriately ambitious goals. The second strategy also involves use of the graphed performance indicators; this time, the teacher uses the graphed database to determine the adequacy of student progress and to decide when programmatic improvements appear warranted.

For both these purposes, the CBM performance indicators are employed. The performance indicators, which provide an overall index of the student's proficiency on the year-long curriculum, are well suited for summarizing the overall rate of student improvement and for making related evaluation decisions, such as judging the appropriateness of the goal and the adequacy of student progress.

Nevertheless, the CBM performance indicators displayed on the student's graph provide relatively little direction for determining the nature of potentially effective program changes. By inspecting the performance indicators to determine the overall rate of growth in the curriculum, the teacher may be able to formulate certain potentially effective instructional changes. For example, with a flat or decelerating slope, hypotheses about (a) the lack of student retention of skills and/or (b) motivation problems can be generated, and related programmatic changes can be considered. However, since the performance indicators do not identify which skills the student currently is performing well and which curricular components the student is not performing proficiently, the practitioner cannot use the performance indicators to formulate decisions about what dimensions of the curriculum might represent an appropriate instructional focus over the next several weeks.

Although the *graphed performance indicators* cannot be used to derive a skills profile on the target curriculum for a given student, the CBM database *does* contain the information required to put together such a skills profile. Since, during CBM testing the student is required to perform skills representing the entire year-long curriculum, student performance on all the curricular content for the year is available for each skill, on any one probe (in math, for example) or across probes (in spelling, for example). Information can be aggregated across probes to formulate a skills analysis of the student's performance.

During the 1987--1988 academic year, Fuchs and associates undertook a series of studies investigating teachers' use of the CBM skills analysis. One study was conducted in math, one in reading, and one two-part study in spelling. The studies all contrasted different types of CBM analyses teachers received to facilitate their instructional decision making. In each study, there was a control group that did not use CBM; a CBM group that relied only on the graphed database, with the related analyses to judge the appropriateness of the goal and the adequacy of student progress; and a CBM group that used both the graphed analyses as well as skills analyses that provided a skills profile to assist the teacher in determining directions for teaching changes. What follows is a detailed description of the methodology, skills analysis procedures, and results for the series of spelling studies, along with a brief description of findings in reading and math.

*Spelling Study 1.* Within the first spelling study, 30 special education teachers were assigned randomly to three groups: control, CBM with graphed analysis, and CBM with graphed analysis plus skills analysis. Each teacher selected two mildly handicapped pupils with spelling IEP goals to participate in the 15-week study. Analyses indicated that teachers and students in the three treatment groups were comparable on demographic variables, including (a) teachers' age, years teaching, years in current position, previous years experience in CBM research projects, highest educational degree, and personal and general teaching efficacy; and (b) students' age, grade, spelling grade level, years in special education, keyboarding skills, handicapping condition, sex, and IQ.

The control teachers in this study implemented their normal procedures for monitoring student progress in spelling. This did not include any use of CBM. As reported by the teachers in posttreatment questionnaires, the control monitoring information primarily consisted of inspection of scores on weekly quizzes assessing student proficiency on weekly spelling lists.

Within the CBM groups, teachers used CBM to monitor their two pupils' progress toward spelling goals. To establish goals, teachers (a) identified the curriculum and the level within the curriculum on which they hoped the student would be proficient by the end of the year, and (b) selected a performance criterion for acceptable performance at the conclusion of the study on April 14.

To monitor student progress toward the performance criterion of the target level of the curriculum, teachers used CBM methodology (Mirkin et al., 1984), in conjunction with computer applications (see Fuchs, Fuchs, & Hamlett, 1988). Each test was created, administered, and scored in the following way. The computer randomly sampled 20 words from the pool of words representing the target level of the spelling curriculum, and printed a hard copy of the 20-word list. A cross-age or peer tutor, aide, or teacher dictated the words from this list, and the student typed the words into the computer, with a maximum of 15 seconds before the computer automatically advanced the student to the next word. If the student finished the word before the 15-second limit, he/she pressed return to advance the computer to the next word. At the end of 20 words or 3 minutes, whichever occurred first, the computer terminated administration of the test and scored the number of correct letter sequences and words. The computer presented these scores to the student, along with a graph showing the numbers of correct letter sequences over time.

Spelling performance was measured in this way at least two times per week. Once each week, teachers used data-management software to inspect the CBM database. This software displayed a graph of the student's number of correct letter sequences over time. This graph also showed (a) broken vertical lines to represent goal changes, (b) solid vertical lines to indicate intervention changes, (c) a "G" to signify the performance criterion expected on April 14, (d) a broken diagonal line to show the goal line, and (e) a solid diagonal line to represent the student's actual rate of progress.

The computer applied the following set of decision rules to the graphs. If the student's actual rate of progress was steeper than the goal line, a decision appeared below the graph saying, "Nice work! Raise your goal." If the student's actual rate of progress was flatter than the goal line, a decision read, "Uh-oh! Make a teaching change." If the student's recent scores were higher than a predetermined ceiling level, a decision read, "Move to the next curriculum level." Finally, if there were fewer than eight new scores since the last vertical line, the decision read, "Insufficient data. Keep collecting data." The computer used an interactive structure to communicate these decisions (see Fuchs, Fuchs, & Hamlett, 1988), where teachers had to inspect the database independently and enter their own decisions. The computer provided corrective feedback to the teachers' responses and provided explanations for correct decisions (see Fuchs, Fuchs, & Hamlett, 1988). CBM teachers in the graphed analysis and in the graphed plus skills analysis received this graphed feedback.

CBM teachers in the graphed plus skills analysis group, however, received additional information. Using the most recent 50 words the student had spelled, the computer provided the following skills analysis. The computer indicated the number of correctly spelled words, the number of Near Misses (incorrect words with at least 50% correct letter sequences), and the number of Far Misses (incorrect words with fewer than 50% correct letter sequences). The computer also identified, for every word in the Near Misses category, the error categories the student had committed, and then showed the teacher (a) for each possible error type, the number of corrects and opportunities, as well as the percentage correct, and (b) three key error categories the student had made most frequently, along with up to four examples of each frequent error category. Finally, the computer presented the teacher with complete lists of the Corrects, Near Misses, and Far Misses. Figures 7 and 8 show a sample 2-page printout of the information contained in the spelling skills analysis.

**Spelling Profile**

| Name: | Domain: Spelling D | Date: 4/15/89 | Page 1 |
|---|---|---|---|

| | |
|---|---|
| Corrects (100 LS Correct): | 14 words(s) |
| Near Misses (60-99% LS Correct): | 19 words(s) |
| Moderate Misses (20-59% LS Correct): | 16 words(s) |
| Far Misses (0-19% LS Correct): | 1 words(s) |

| Type | Correct | Possible | PCT |
|---|---|---|---|
| FSLZ | 0 | 0 | – |
| Final E | 1 | 5 | 20 |
| Blend | 7 | 10 | 70 |
| Double | 3 | 4 | 75 |
| Dual Con | 12 | 24 | 50 |
| Vowel+R | 9 | 14 | 64 |
| Vowel+N | 6 | 8 | 75 |
| Suffix | 4 | 6 | 67 |
| Digraph | 7 | 10 | 70 |
| Vow Team | 4 | 12 | 33 |
| C/S | 0 | 1 | 0 |
| -Le Word | 4 | 7 | 57 |
| Final Vow | 3 | 7 | 43 |
| Ild Word | 0 | 0 | – |
| Dge Word | 0 | 1 | 0 |
| Ch/tch | 2 | 2 | 100 |
| C/ck | 0 | 2 | 0 |
| Shun Word | 0 | 1 | 0 |
| Combo | 1 | 1 | 100 |
| Ign/igh | 0 | 0 | – |
| V+L+Con | 0 | 0 | – |
| Sure Word | 0 | 0 | – |
| Ance Word | 0 | 0 | – |
| Irregular | 1 | 1 | 100 |
| Apos'Phe | 0 | 0 | – |
| Sing Vow | 21 | 30 | 70 |
| Sing Cons | 47 | 49 | 96 |

**Key Errors**

| Vow Team | Dual Con | Final E |
|---|---|---|
| Instead-Insted | Learner-Leaner | Alone-Alon |
| Moisten-Mosten | Sample-Samble | Knife-Knif |
| Quieter-Quiter | Chart-Chard | Rare-Rar |
| Trouble-Trubble | Mumble-Mobble | Cube-Cub |
| Rail-Real | Tractor-Trater | |
| Certain-Chanten | Apart-Apcot | |

Figure 7. Page 1 of the computerized CBM spelling skills analysis.

---------------------- Moderate Misses (20-59% LS Correct) ----------------------

| | | | |
|---|---|---|---|
| 57 Tickle-Teakle | C/CK | Sing Vow | |
| 57 French-Fanch | Vowel + N | Blend | |
| 57 Mumble-Mobble | Dual Con | Sing Vow | |
| 50 Unlucky-Unluke | Final Vow | C/CK | |
| 50 Tractor-Trater | Vowel + R | Dual Con | |
| 50 Apart-Apcot | Vowel + R | Dual Con | |
| 44 Calendar-Cander | Vowel + R | Vowel + N | Sing Cons |
| 42 Mumble-Mommbe | -Le Word | Sing Vow | |
| 40 Rail-Real | Vow Team | | |
| 37 Station-Stanch | Shun Word | | |
| 28 Sample-Scembe | -Le Word | Dual Con | Sing Vow |
| 25 Certain-Chanten | Vow Team | Vowel +R | C/S |
| 25 Squeeze-Sccase | Vow Team | Digraph | Digraph |
| 20 Limb-Lem | Dual Con | Sing Vow | |
| 20 Treatment-Tempemt | Suffix | Vow Team | Blend |
| 20 Limb-Leam | | Dual Con | Sing Vow |

---------------------- Far Misses (0-19% LS Correct) ----------------------

| | | | |
|---|---|---|---|
| 14 Giggle-Gelly | -Le Word | Double | Sing Vow |

Figure 8. Page 2 of the computerized CBM spelling analysis.

Several types of outcome measures were collected. First, fidelity of treatment was indexed. Second, teachers' program development was measured in several ways. Finally, student achievement was assessed using a standardized spelling achievement test, which required students to write Grades 1-6 words that appear with high frequency across curricula. Results indicated the following.

With respect to fidelity of treatment, teachers in the two CBM groups structured their measurement procedures and actually measured student performance in a highly accurate and comparable manner. However, teachers in the graphed-plus-skills-analysis group received relatively high fidelity of treatment scores for the Evaluation component of the fidelity of treatment scale; their Instructional Plan Sheets, on which they recorded their teaching changes, were completed in a more acceptable fashion, compared to the graphed-analysis-only teachers.

In a related way, for program development, teachers in the two CBM groups scored comparably on most variables, including number of goal increases, level of goal ambitiousness, and number of teaching changes. However, teachers in the graphed-plus-skills-analysis group received higher scores than teachers in the graphed-analysis-only group on the number of skills they targeted for instruction and listed on their Instructional Plan Sheets.

In terms of achievement, teachers in the graphed-plus-skills-analysis group effected greater growth compared to (a) teachers in the graphed-analysis-only group and (b) teachers in the control group. The average gains from pre- to posttesting for the graphed-plus-skills-analysis group, the graphed-analysis-only group, and the control group, respectively, were approximately 37, 14, and 12.

Consequently, it appeared that the skills analysis information contributed critical information in order to promote effective instructional planning. With the addition of the skills analysis to the graphed feedback, teachers were able to write more acceptable instructional programs; they cited more skills to target during instruction; and they effected superior student achievement. Results of this study strongly support the usefulness of skills analysis within CBM to support teachers' effective instructional decision making.

Nevertheless, an important shortcoming of this study, with respect to generalization to typical CBM procedures, is that the graphed-analysis-only procedures used in this study involved computerized data collection. This meant that teachers did not routinely inspect students' spelling performance. Yet, with typical CBM, which does not rely on automatic data collection, teachers frequently score and thereby inspect student spelling samples. With computerized data collection, however, teachers do not routinely score student tests. Rather, they typically see *only* the graphed analysis. Because of this limitation associated with the computerized data collection used in this study, a second, related investigation was undertaken. (For a complete description of this study, see Fuchs, Fuchs, Hamlett, & Allinder, in press.)

*Spelling Study* 2. In this second study, the 30 same teachers were assigned randomly to three treatment groups: control, CBM with graphed-plus-skills analysis, and CBM with graphed analysis plus Near Misses inspection. Study procedures were identical to those employed in Study 1, with the following deviation. This time, CBM teachers who did not receive the skills analysis did have the opportunity to inspect student spellings. This was accomplished in the following way. After viewing graphs and receiving the graphed analysis, teachers in the graphed analysis plus Near Misses inspection group saw the list of Near Misses. The Near Misses list contained incorrectly spelled words from the pool of the most recent 50 words the student had spelled on his/her tests. These Near Misses had to be at least 50% correctly spelled, in terms of letter sequences. They were presented to the teachers from most correct (99% letter sequences correct) to least correct (50% letter sequences correct), with the correct and incorrect spellings next to each other. (See page 2 Near Misses of Figure 8; however, only the correct and incorrect spelling were provided in this Near Misses treatment.)

This Near Misses condition was incorporated into Study 2 in order to provide teachers, who did not receive formal skills analysis, an opportunity to view a structured presentation of the student's spelling errors. This structuring of the student's Near Misses provided richer information than the graphed analysis only condition of Study 1 and therefore better approximated typical CBM procedures where teachers score student tests by hand. Nevertheless, the Near Misses condition provides a more systematic and structured presentation of information than is inherent within the simple hand scoring teachers complete with noncomputerized CBM. Consequently, the Near Misses condition must be viewed as a form of CBM that presents teachers with information somewhat less organized than skills analysis but more systematic than provided by simple hand scoring.

Results of this second study indicated the following. CBM teacher performance was comparable on fidelity of treatment and program development indices. However, teachers did effect differential achievement among their students. Progress for the students within the graphed-plus-skills-analysis groups was reliably better than that of controls (an average gain of approximately 33 versus approximately 12). However, the difference in achievement between the Near Misses group and the control-only group approached statistical significance ($p = .07$), with mean gains of approximately 24 versus 12. The difference in growth between the skills analysis and the Near Misses group was not reliably different. (For a complete description of this study, see Fuchs, Allinder, Hamlett, & Fuchs, in press.)

This series of studies suggests the following. First, skills analysis does seem to provide teachers with structured information that supplements the graphed CBM database in such a way that facilitates teachers' effective instructional decision making. Second, as additional sources of structured feedback are provided to teachers (graphed analysis vs. Near Misses lists vs. skills analysis), teachers' instructional decision making and student achievement appears to be enhanced.

*Reading and math studies.* During the 1987-1988 academic year, similar studies were conducted in the areas of reading and math. In these additional academic areas, CBM teachers either received graphed feedback only or graphed feedback with skills analysis. In both additional academic areas, results were similar to those found in spelling. That is, with the additional information supplied by the skills analysis, teachers were able to structure better instructional programs and they effected superior student achievement. Consequently, the finding that teachers can use additional sources of feedback about student performance, including skills analysis, to enhance instructional decision making appears to be robust. (For descriptions of the reading and math studies, respectively, see Fuchs, Fuchs, & Hamlett, 1989 and Fuchs, Fuchs, Hamlett, & Stecker, 1990.)

## Concluding Remarks: Getting Teachers to Use CBM

This review of research highlights three ways in which teachers may use CBM databases to assist in their instructional decision making: (a) to monitor the appropriateness of their goals and to adjust goals as necessary, (b) to judge the adequacy of student progress and to create instructional modifications when needed, and (c) to rely on skills analysis to derive additional information from the CBM database for formulating potentially effective instructional improvements.

As noted, studies have documented that CBM can be used to effect statistically significant and practically important differences in student achievement outcomes across academic areas. Yet, as noted by Wesson, King, and Deno (1984) and others (e.g., Walton, 1986), teachers are reluctant to employ CBM and other forms of ongoing student performance monitoring, because these measurement systems are time consuming and frequently technically demanding (see Wesson, Fuchs, Tindal, Mirkin, & Deno, 1986).

A pressing question, then, is: How can we facilitate teachers' implementation of ongoing assessment systems and induce teachers to use these systems effectively? Our CBM intervention research suggests the following. First, computers can be used to reduce teacher time necessary to implement CBM. With computerized automatic data collection in reading, spelling, and math (Fuchs et al., 1990), the teacher is freed from the time-consuming tasks of developing measures, administering and scoring tests, and analyzing student performance profiles. Rather, once students have been taught to use the CBM software, teachers need only to view assessment profiles (i.e., graphs and skills profiles that are produced automatically by computers). Evidence indicates that with these automatic data collection and analysis programs, teacher time devoted to measurement can be virtually eliminated and teacher satisfaction with CBM improves (Fuchs, Hamlett, Fuchs, Stecker, & Ferguson, 1988).

Despite this improved feasibility, it appears that teachers may still require some inducement to incorporate the information presented in CBM assessments into their instructional decision making. Research (e.g., Fuchs, Fuchs, Hamlett, & Ferguson, 1989; Tindal, Fuchs, Mirkin, Christenson, & Deno, 1981) indicates that teachers may experience difficulty in formulating effective strategies for revising their instruction when student performance data indicate that student rates of progress are inadequate. Additionally, given the increasing numbers of students on many special education roles and the complexity and diversity of class compositions in regular and special education settings, the individual nature of the CBM assessment profiles and instructional implications may be problematic for teachers. That is, teachers may recognize not only *when* they need to revise different students' programs, but also *how* they might improve student programs. Yet, the numbers and types of students and the many *different* instructional adaptations indicated by the CBM data may preclude or reduce the likelihood of teachers' responsive use of a CBM database.

In our CBM research we have tried to address these two problems (i.e., teachers' need for assistance in formulating potentially effective revisions to their students' instructional programs and the logistical difficulties in revising different students' programs in different ways at different times), in several ways. First, in terms of support to teachers in order to assist them in formulating potentially effective instructional revisions, consultants (i.e., our project staff) visit teachers once every 1-2 weeks, review with them the CBM student profiles, and assist them in identifying instructional revisions, including the provision of instructional packets to assist teachers in specifying and implementing instructional modifications.

Second, as a alternative to frequent consultant visits, we have developed and researched computerized expert systems that provide systematic consultation in reading, spelling, and math. Our initial research (e.g., Fuchs, Fuchs, Hamlett, & Ferguson, in press) using these computerized recommendation systems indicates that they may represent an effective substitute for the relatively expensive use of consultants.

Third, with respect to the logistical problems of implementing many programmatic changes for different students at different times, we have begun to develop and research computer programs that simultaneously consider all students on an individual teacher's caseload. These programs present information and make instructional suggestions for flexible groupings of students, rather than for individuals. We hope that with these group profiles and recommendations, teachers will revise instructional groupings more frequently and implement sound instructional strategies for these flexible student groupings. Research investigating this possibility is under way.

## REFERENCES

Anderson-Inman, L., Walker, H., & Purcell, J. (1984). Promoting the transfer of skills across settings: Transenvironmental programming for handicapped students in the mainstream. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp.17-37). Columbus, OH: Merrill.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122-147.

Blankenship, C. (1985). Using curriculum-based assessment to make instructional decisions. *Exceptional Children, 52*, 233-238.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children, 19* (8), 1-16.

Fuchs, L. S., Allinder, R. M., Hamlett, C. L., & Fuchs, D. (1990). An analysis of spelling curricula and teachers' skills in identifying error types. *Remedial and Special Education, 11* (1), 42-53.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal, 21,* 449-460.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation on student achievement: A meta-analysis. *Exceptional Children, 53,* 199-208.

Fuchs, L. S., & Fuchs, D. (1990). Curriculum-based assessment. In C. Reynolds & R. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children (Vol. 1): Intelligence and achievement* (pp. 435-455). New York: Guilford Press.

Fuchs, L. S., Fuchs, D., & Deno, S. L. (1985). The importance of goal ambitiousness and goal mastery to student achievement. *Exceptional Children, 52,* 63-71.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1988). Computers and curriculum-based measurement: Effects of teacher feedback systems. *School Psychology Review, 18,* 110-123.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989a). Effects of alternative goal structures within curriculum-based measurement. *Exceptional Children, 55,* 429-438.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989b) . Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10* (2), 43-52.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L., (in press). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *Journal of Educational Research.*

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. (in press). The contribution of skills analysis to curriculum-based measurement in spelling. *Exceptional Children.*

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (in press). Effects of expert system advice within curriculum-based measurement using a reading maze task. *Exceptional Children.*

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review, 19,* 6-22.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1990). *Basic Reading, Basic Spelling, Basic Math.* [Computer Programs.] Austin: PRO-ED, Inc.

Fuchs, L. S., Hamlett, C. L., Fuchs, D., Stecker, P., & Ferguson, C. (1988). Conducting curriculum-based measurement with computerized data collection: Effects on efficiency and teacher satisfaction. *Journal of Special Education Technology, 9* (2), 73-86.

Gagne, R. M. (1964). The implications of instructional objectives for learning. In C. M. Lindvall (Ed.), *Defining educational objectives.* Pittsburgh: University of Pittsburgh Press.

Gickling, E. E., & Thompson, V. (1985). A personal view of curriculum-based assessment. *Exceptional Children, 52,* 205-218.

Goodstein, H. A. (1982). The reliability of criterion-referenced tests and special education: Assumed versus demonstrated. *The Journal of Special Education, 16,* 37-48.

Hartley, J., & Davies, I. K. (1976). Preinstructional strategies: The role of pretests, behavioral objectives, overviews, and advance organizers. *Review of Educational Research, 46,* 239-265.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. (1978). *Measuring pupil progress toward the least restrictive environment* (Monograph No. 10). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities. (ERIC Document Reproduction Service No. ED 185 767)

Mager, R. F. (1975). *Preparing instructional objectives.* Belmont, CA: Fearon Publishers.

Marston, D., Fuchs, L. S., & Deno, S. L. (1985). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique, 11,* 77-90.

McNeil, J. D. (1967). Concomitants of using behavioral objectives in the assessment of teacher effectiveness. *Journal of Experimental Education, 35,* 69-74.

Mirkin, P. K., Deno, S. L., Fuchs, L. S., Wesson, C., Tindal, G., Marston, D., & Kuehnle, K. (1984). *Procedures to develop and monitor progress toward IEP goals* (2nd ed.). Minneapolis: University of Minnesota.

Peckham, P. D., & Roe, M. D. (1977). The effects of frequent testing. *Journal of Research and Development in Education, 10* (3), 40-50.

Rosswork, S. G. (1977). Goal setting: The effects on academic tasks with varying magnitudes of incentive. *Journal of Educational Psychology, 69,* 710-715.

Salvia, J., & Hughes, C. (1990). *An introduction to the assessment of achievement.* New York: Macmillan.

Shinn, M. R. (Ed.) (1989). *Curriculum-based measurement: Assessing special children.* New York: Guilford Press.

Shinn, M. R., Rosenfield, S., & Knutson, N. (1989). Curriculum-based assessment: A comparison of models. *School Psychology Review, 18,* 299-316.

Tindal, G., Fuchs, L. S., Fuchs, D., Shinn, M. R., Deno, S. L., & Germann, G. (1985). Empirical validation of criterion-referenced tests. *Journal of Educational Research, 78*, 203-209.

Tindal, G., Fuchs, L. S., Mirkin, P. K., Christenson, S., & Deno, S. L. (1981). *The effect of measurement procedures on student achievement* (Research Report No. 61). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities. (ERIC Document Reproduction Service No. ED 218 847)

Tucker, J. (1987). Curriculum-based assessment is no fad. *The Collaborative Educator, 1*(4), 4, 10.

Walton, W. T. (1986). Educators' responses to methods of collecting, storing, and analyzing behavioral data. *Journal of Special Education Technology, 7*, 50-55.

Wesson, C., Fuchs, L. S., Tindal, G., Mirkin, P. K., & Deno, S. L. (1986). Facilitating the efficiency of on-going curriculum-based measurement. *Teacher Education and Special Education, 9*, 166-172.

Wesson, C., King, R., & Deno, S. L. (1984). Direct and frequent measurement: If it's so good for us, why don't we do it? *Learning Disability Quarterly, 7* (1), 45-48.

White, O. R. (1984). Descriptive analysis of extant research literature concerning skill generalization and the severely/profoundly handicapped. In M. Boer (Ed.), *Investigating the problem of skill generalization: Literature review* (pp. 1-19). Seattle: University of Washington, Washington Research Organization.

White, O. R., & Haring, N. G. (1980). *Exceptional teaching* (2nd ed.). Columbus, OH: Merrill.

Yalow, E. S., & Popham, W. J. (1983). Content validity at the crossroads. *Educational Researcher, 12* (8), 10-14, 21.